AJPS AMERICAN JOURNAL of POLITICAL SCIENCE

The Disclosure Dilemma: Nuclear Intelligence and International Organizations

Allison Carnegie Columbia University
Austin Carson University of Chicago

Abstract: *Scholars have long argued that international organizations solve information problems through increased transparency. This article introduces a distinct problem that instead requires such institutions to keep information secret. We argue that states often seek to reveal intelligence about other states' violations of international rules and laws but are deterred by concerns about revealing the sources and methods used to collect it. Properly equipped international organizations, however, can mitigate these dilemmas by analyzing and acting on sensitive information while protecting it from wide dissemination. Using new data on intelligence disclosures to the International Atomic Energy Agency and an analysis of the full universe of nuclear proliferation cases, we demonstrate that strengthening the agency's intelligence protection capabilities led to greater intelligence sharing and fewer suspected nuclear facilities. However, our theory suggests that this solution gives informed states a subtle form of influence and is in tension with the normative goal of international transparency.*

Replication Materials: The data required to verify the results in this article are available on the *American Journal of Political Science* Dataverse within the Harvard Dataverse Network, at: https://doi.org/10.7910/DVN/JQEUBQ.

Theories of international cooperation have long argued that states follow international rules because noncompliance can prompt naming and shaming, damage reputations, and lead to sanctions or military action. To punish states effectively, the international community must first learn about such breaches, yet this is difficult since states hide their violations and international organizations often lack the authority and resources to intrusively monitor compliance. While many states possess extensive intelligence capabilities that allow them to detect violations of treaties and norms, they often refuse to fill these informational gaps. This article develops a pervasive reason for this reticence: States fear the exposure of sensitive details about their intelligence collection process, since this can jeopardize their ability to obtain such information in the future. It also develops a surprising solution—building new secrecy capabilities into international organizations — that reduces the damage from information disclosures but also enables a subtle method of exerting power.

The challenges of monitoring nuclear proliferation illustrate this broad problem. In 1978 and 1979, American intelligence discovered new signs of progress in Pakistan's uranium enrichment and plutonium reprocessing efforts (Richelson 2007, 339). Islamabad's program posed a threat to both regional stability and the non-proliferation regime, and the United States wrestled with how to address it. One option was to share its insights with other states, which might help to close off procurement routes and improve the responsiveness of key international audiences. However, American leaders feared leaks. The State Department cautioned that it "[did]

Allison Carnegie is Assistant Professor, Department of Political Science, Columbia University, 1331 IAB 420 W. 118th Street, New York, NY 10027 (allison.carnegie@columbia.edu). Austin Carson is Assistant Professor, Department of Political Science, University of Chicago, Pick Hall 428, 5828 S. University Avenue, Chicago, IL 60637 (acarson@uchicago.edu).

We thank Jeff Colgan, Sarah Daly, Christina Davis, Lindsay Dolan, Page Fortna, Andrew Gelman, Thomas Groll, Kolby Hanson, Alisha Holland, Turkuler Isiksel, Ashley Jester, Leslie Johns, Jeff Kaplow, Andrew Kerner, Chris Kojm, Austin Long, Gwyneth McClendon, John Marshall, Krzysztof Pelc, Maggie Peters, Paul Poast, Carlo Prato, Tonya Putnam, David Steinberg, Joshua Simon, Arthur Spirling, Felicity Vabulas, Jane Vaynman, and Tristan Volpe for helpful comments. We also thank the participants of the APSA 2016 and 2017 and IPES 2015 conferences and seminars at Columbia University, Georgetown University, George Washington University, Harvard University, Rutgers University, Stanford University, Temple University, the University of Chicago, the University of Michigan, and the University of Pennsylvania. We received excellent research assistance from Justin Canfil, Donald Casler, Richard Clark, Andrew Kosenko, Manuel Puente, Bryan Schonfeld, Laura Sipe, and Joon Yang. We gratefully acknowledge funding from the Columbia Provost's Grant. All remaining errors are our own.

American Journal of Political Science, Vol. 63, No. 2, April 2019, Pp. 269–285

DOI: 10.1111/ajps.12426

not want to contact" governments such as India, Israel, and Taiwan because "they could threaten the effectiveness of our efforts by informing the [Government of Pakistan], or by making our efforts publicly known."[1] As a result, these intelligence-driven findings were not widely shared.

We refer to this situation as a "disclosure dilemma," which arises when states possess private information whose wide dissemination would yield political benefits but would also trigger negative political or operational externalities. Although several kinds of information can prompt disclosure dilemmas, we focus on a particularly widespread and consequential type: intelligence. Revealing intelligence can further a state's political goals—especially regarding unfriendly actors like adversaries and terror groups—by, for example, facilitating the punishment of war crimes, galvanizing the scrutiny of clandestine nuclear programs, or improving the operational security of peacekeeping missions. However, doing so can also tip off current and future intelligence targets.[2] For example, when Iraq found out that its nuclear activities had been tracked via intercepted cell phone conversations, it switched to alternative communication methods.[3]

However, without intelligence details, observers find it difficult to discern the truth. After all, informed states have political incentives to lie to discredit their adversaries and secure international support for their policy goals.[4] Yet while such dilemmas might appear intractable, we argue that international organizations (IOs) can often mitigate them. When IOs are equipped to receive and protect sensitive information, they encourage states to share it by minimizing the threats to states' future intelligence collection. Moreover, IOs can use their technical expertise and monitoring abilities to vet intelligence-based claims, solving persistent credibility problems that occur if informed states make assertions without disclosing evidence. Easing these dilemmas using IOs can result in a greater amount and higher quality of available information, which, in turn, can improve international monitoring efforts and overall compliance with a given regime.

While our framework helps to make sense of IOs that protect sensitive information in a variety of domains, including international war crimes, international trade and investment, and peacekeeping, we evaluate these claims in the nuclear nonproliferation arena for several reasons. Theoretically, the nuclear domain is a high-stakes area of considerable scholarly interest. It also represents a tough case for our theory due to the large cost associated with revealing intelligence details about detected nuclear activities. Empirically, we take advantage of newly available archival materials and changes to the International Atomic Energy Agency's (IAEA) intelligence safeguarding capabilities. Around 1991, the IAEA embraced a new role for shared intelligence in reaction to the end of the Cold War and the revelation of covert Iraqi nuclear facilities after the Persian Gulf War. We collect new data on American intelligence disclosures from archival documents, interviews with current and former officials, and secondary sources[5] to assess differences before and after these reforms as well as their impact on proliferation activities. We then conduct case study analyses that draw on the universe of cases of nuclear proliferation since 1970, finding that the IAEA's reforms successfully elicited more robust intelligence disclosures and improved the scrutiny of hidden nuclear facilities in countries that were non-allies of the United States.

We thus document states' unexpected willingness to share sensitive intelligence details with an IO, and describe improvements in the IO's performance due to greater institutional secrecy rather than transparency. In doing so, we identify a novel function for IOs that is broadly applicable and modifies the literature's central claim that institutions facilitate cooperation by making information widely and equitably available (e.g., Keohane 1984, 94). This conventional view suggests that the broad dissemination of compliance-related information to both states and non-state actors can enable political processes and pressures that improve compliance. We extend and revise this account, arguing that *sensitive* compliance information like intelligence requires IOs to protect—rather than widely disseminate—key details provided by states in order to monitor compliance effectively. To solve disclosure dilemmas, IOs must therefore develop a secrecy capacity, which can conflict with the international community's deepening expectations of international transparency

[1] "Pakistan Proliferation Problem," State to U.S. Embassy United Kingdom, Cable 292469, November 18, 1978. Wilson Center Digital Archive.

[2] Carnegie and Carson (2018) show that revealing intelligence can cause others to violate the rules by sharpening threat perceptions and increasing pessimism about compliance, and do not assess a role for international organizations. Other interesting work demonstrates that providing information about states' own capabilities can threaten their safety during peacetime (Coe and Vaynman 2018) and damage their battlefield performance during war (Lindsey 2015; Slantchev 2010). See also Carnegie and Carson 2019.

[3] See the supporting information.

[4] For instance, the United States and the Soviet Union frequently lied about each others' activities during the Cold War. See, for example, "New Pieces in the Puzzle of Flight 007," *The Nation*, August 17, 1985. See also the Iraq case study in the supporting information.

[5] Because our interviewees requested anonymity, we refer to each by number.

and accountability. Moreover, we demonstrate that addressing disclosure dilemmas endows some states with a subtle form of power, enabling them to quietly turn the intelligence supply spigot on or off depending on whether the country in question is an ally or a rival.

Information, Institutions, and Compliance

Scholars have long recognized that the underprovision and uneven distribution of information can allow defections from agreements to go undetected, breeding fears of exploitation and preventing states from cooperating.[6] Improving transparency gives states and other actors the raw material to scrutinize noncompliance and punish it, making "compliance information to facilitate compliance with international agreements . . . a centerpiece of neoliberal institutionalism" (Dai 2002, 409). We focus on a particular kind of compliance information—national intelligence — that is information gathered through clandestine means about other states' behavior and intentions (Warner 2014, 21). Most states have systematized bureaucracies for intelligence collection and draw on human and technical intelligence collection methods including spy rings, imagery intelligence (i.e., satellite photography), signals intelligence (i.e., interception of cellular or other means of communication), and more obscure tactics (i.e., nuclear radiation emissions analysis; Richelson 2007). Intelligence can be highly germane to questions of compliance either because states seek information about such behavior or because they obtain it as a by-product of analyzing other intelligence priorities. Informed states may be tempted to disclose this information in order to advance their practical, diplomatic, and strategic goals, including to facilitate the broad scrutiny and punishment of breaches of international norms and laws (Chesterman 2006).

Yet while sharing intelligence can facilitate cooperation in accordance with the standard view, we show that because such information is *sensitive*, transparency is not an unmitigated good. Disclosing intelligence often has an important downside: It endangers the deception necessary for future intelligence collection. For instance, sharing intelligence from signals intercepts can lead a target to change its use of a vulnerable cell phone; providing imagery might give current or future targets tips on how to avoid observation; and sharing intelligence from a human source can lead to that source's expulsion, imprisonment, or death (Richelson 2007). States may try to avoid these issues by sharing their conclusions rather than the details of their sensitive sources and methods, but this creates a credibility problem. Because intelligence sharers often have well-known political interests at stake, states view unilateral intelligence claims without details with suspicion. An informed state could lie in order to, for example, justify sanctions or a military intervention against an adversary, or even to influence international opinion about such a state.[7] Alternatively, a state could only reveal its sources to trusted allies (Walsh 2010), but this substantially limits the practical and political impact. Most state and non-state audiences remain in the dark in this scenario, which hinders the likelihood, strength, and durability of any multilateral response. Thus, we show that these dilemmas often lead states to withhold intelligence from international and domestic audiences, which damages international cooperation efforts.

IOs and Disclosure Dilemmas

We argue that international organizations constitute a potential solution to disclosure dilemmas. IOs can apply their technical expertise and their own information to detect states' mistaken or manipulated claims, boosting the information's political impact and increasing support for punitive action against a violator.[8] Moreover, an IO with monitoring powers can follow up on intelligence-based "tips" by asking additional questions and scrutinizing previously hidden or unreported activities.[9] Potential violators may allow an IO—but not a state—to do so because they may view an IO as less biased or politically motivated. If an IO can also protect sources and methods details, informed states may be willing to disclose them to the IO. Importantly, IOs themselves rarely possess a standing intelligence collection capability because states hesitate to yield informational sovereignty or to pay the associated steep financial costs (Carment and Rudner 2007, 2).

We claim that an IO can serve this purpose under two primary conditions. First, an IO must have the capacity to credibly review, assess, and act on sensitive information,

[6] For example, see Keohane (1984).

[7] For instance, the United States exaggerated its intelligence about Iraq for this reason in 2003, as we detail in the case studies.

[8] On vetting claims based on sensitive information, see Chesterman (2006, viii). On IOs' ability to validate policy proposals in a relatively unbiased manner, see Voeten (2005) and Thompson (2006).

[9] Note that only one of these functions (monitoring) requires the IO to have a close relationship with the potential violator. Using relatively unbiased technical experts to vet intelligence-based claims is broadly applicable.

which requires a reputation for technical expertise and relatively unbiased judgment; otherwise, it cannot provide added legitimacy to address the intelligence holder's credibility problem. Second, it must be designed to receive and protect sensitive information by limiting its dissemination within the IO and preventing unauthorized leaks. Without confidentiality protections, all of the information that an IO receives or collects is widely disseminated within the IO secretariat, among member states, and often in publicly accessible formats. In contrast, IOs with an effective secrecy capacity—such as secure cyber and physical storage, document classification schemes, and staff policies for information disclosure—have procedures in place to limit the dissemination of certain kinds of information. The better an IO can keep sensitive details secret, the more states share such information with it.[10]

However, states may not reveal information about all intelligence targets equally; just because states *can* reveal information without jeopardizing their intelligence collection capabilities does not mean that they *will.* In particular, states holding intelligence-based insights should prioritize disclosing information about non-allies and refuse to disclose intelligence that negatively implicates allies, since doing so could threaten their cooperative relationships, jeopardize political goals, and weaken states that are important for their national security. Moreover, multilateral cooperation is often less important for allies because intelligence holders can exert direct influence over their friends, with whom they often have extensive trade, financial, and other ties (Miller 2014). We therefore expect a state with intelligence about violations of rules or norms to disclose it to an IO when (1) the IO is equipped to receive and protect it and (2) the state's intelligence does not implicate its allies.

A Model of Intelligence Disclosures

We develop a simple formal model as a heuristic device to understand the cooperation problems that emerge when compliance information is sensitive, and to examine how they affect intelligence sharing and proliferation. The model makes our assumptions transparent, clarifies how our argument differs from extant scholarship, and allows us to rigorously derive testable implications. This section describes the model in brief; the proofs appear in the supporting information due to space constraints.

The game features three actors: state E (an intelligence holder), state A (a potential rule violator), and I (the international community). State A can be of two types: $A \in \{A^H; A^L\}$. A^H has high incentives to violate international law, whereas A^L has low incentives to do so. The prior probability on state A's type is $prob(A = A^H) = \theta < 1/2$, which is common knowledge.

The game begins with nature choosing A's type. A observes its type and chooses whether to violate international law ($v = 1$) or not ($v = 0$). After observing A's action and type,[11] E decides whether to publicly declare that A violated the law ($x_E = D$) or not ($x_E = \neg D$). If A violated the law, E can also decide whether to reveal its sources and methods, which prove to I that the violation occurred ($x_E = DM$).[12] After I observes E's action, but not A's action or type, it chooses whether to sanction A, represented by $e \in \{0, 1\}$.

The actors' utilities are determined as follows. A receives both benefits and costs from violating the law, summarized by the term v. The benefits outweigh the costs when it is type A^H, whereas the costs dominate when it is type A^L. It also loses utility from any enforcement efforts by I, where the term α captures how much utility A loses from these efforts. This could reflect A's attempts to resist sanctions, its dependence on I for investment or other resources, or other factors. A's utility function is thus $u_{A^H} = v - \alpha e$ for the high type and $u_{A^L} = -v - \alpha e$ for the low type.

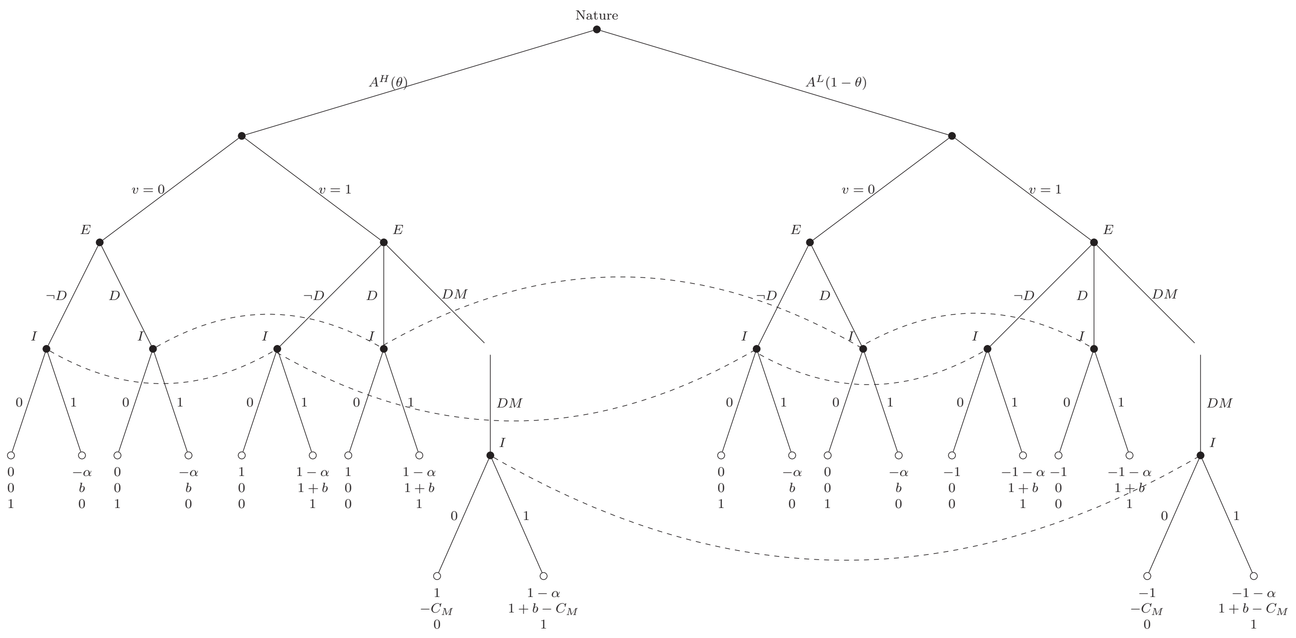
State I wants to enforce the law in an unbiased manner such that its utility function is $u_I = \mathbb{1}\{e = v\}$, where $\mathbb{1}\{\}$ is an indicator function. However, E has political considerations that lead to bias b against A. E's bias prevents it from desiring perfect enforcement, as it may prefer punishment against A even if no violation occurred. For example, A could be an adversary, and E could seek to weaken it for political and economic reasons. E also loses utility from several activities: First, E pays C_M for revealing sources and methods due to the damage that it causes

[10] Considerable variation exists in the degree to which IOs can protect sensitive information in practice. Some IOs do not do so because of political impediments such as lack of political support, status quo bias, or sovereignty concerns. Others face economic constraints due to the cost of implementing reforms, or they do not encounter sensitive information. We investigated patterns across 106 IOs and found that 51 of them deal with sensitive information. Thirty-seven IOs have special protocols in place, 18 use special document classification schemes, 19 feature specific staff policies to handle such information, and 23 have physical or cyber storage requirements. We provide more details in the supporting information.

[11] We assume that E observes A's type for simplicity, and due to its sophisticated intelligence collection capabilities.

[12] Sources and methods cannot be revealed if A did not violate the law since we assume that E cannot falsely provide this proof. This assumption could be easily relaxed.

FIGURE 1 Game Tree



to future intelligence collection.[13] E's utility is given by $u_E = (v + b)e - C_M\mathbb{1}\{x_E = \{DM\}\}$. These payoffs are depicted in the game tree (Figure 1).

A weak perfect Bayesian Nash equilibrium is a tuple $\Psi = ((v_{A^H}, v_{A^L}); x_E(\cdot); e(\cdot); p)$ where

- $v_A \in \{0; 1\}$, $A = \{A^H; A^L\}$.
- $x_E : \{0; 1\} \to \{\neg D; D; DM\}$ represents E's response to A's action v.[14]
- $e : \{\neg D; D; DM\} \to \{0; 1\}$ represents I's response to E's action.
- $p = \{p_{\neg D}; p_D; p_{DM}\} \in \{0; 1\}^3$ are the probabilities that I assigns to $v_A = 1$ given E's action.
- Strategies are optimal given beliefs and beliefs satisfy Bayes' Rule given strategies, whenever possible.

We make three main assumptions. First, we assume that $\theta < 1/2$, so that in the absence of any information, A is more likely to have low incentives to break the law. Thus, I does not punish A unless E provides sufficient proof that A violated the law. Second, $\alpha > 1$, so that A does not want to violate the law if I would then punish it.[15] Third, $C_M > 1$ without an IO.

Solution

The equilibria depend on the value of E's bias against A. When the bias is small, or $b < -1$, no disclosure dilemma exists because E does not want to punish A for a violation. If E were to disclose its conclusions, I might sanction A, but E does not, so A's violation goes unpunished. This could occur, for example, if E is friendly with A, so that harming A could hurt E as well by threatening its security or economic interests. At the other extreme, when E's bias is large ($b > +C_M - 1$), no disclosure dilemmas exist either because E always discloses its sources and methods. E so strongly wants I to punish A that it provides all of its sensitive details, and, knowing that this will occur, A does not break the law.

The disclosure dilemma bites, however, for intermediate values of E's bias ($-1 < b < C_M - 1$). When the bias is in the lower end of this range ($b < 0$), E's conclusions alone might be credible, and I punishes A after E merely states that a violation occurred. However, another equilibrium exists in which I disbelieves E's accusation unless E provides sources and methods. Since doing so

[13] This represents in reduced form the intuition of a dynamic model in which A could adapt to the revelation of E's sources and methods to avoid detection. It would be straightforward to add a subgame in which A could violate again, after which E would have a reduced chance of detecting a violation if it revealed sources and methods in the previous period absent an IO.

[14] In principle, x_E depends on A's type and action, but since E's utility depends only on A's and I's actions, in equilibrium x_E only depends on A's action.

[15] This ensures that the parameter space exists in which E's claims that lack sources and methods are not credible, but revealing its sources is too costly.

would compromise E's future intelligence collection capabilities, it does not, and A's violations are unidentified, unpunished, and therefore undeterred. If E's bias is in the higher end of this range ($b > 0$), the first equilibrium disappears, as it is no longer credible that if I responds to E's accusations by punishing A, E does not accuse A of a violation when A is innocent.

Suppose now that the states play the same game, but E has the option of securely disclosing intelligence—including sources and methods details—to an IO. We assume that the IO lowers C_M enough so that $b > C_M - 1$. As explained in the previous section, an IO can validate and/or supplement E's intelligence and then disseminate its conclusions to I. Further, because the IO lacks E's political bias, I tends to believe the IO's conclusions. In this case, the equilibrium disappears in which I does not punish unless E reveals sources and methods, but E does not do so due to the cost. We are left with two possible equilibria. Either I believes E's accusations, or it does not but E is willing to reveal its sources to trigger I's punishment. In either case, A chooses not to break the law. Figure 2 demonstrates this visually, showing how the equilibria change after an IO is introduced that reduces the value of C_M. For illustrative purposes, we depict the extreme case where $C_M = 0$.

Of course, if E's bias is small enough, lowering C_M still does not lead E to disclose its intelligence to the IO. Put differently, a properly equipped IO encourages E to disclose intelligence, including sources and methods, only if E has political incentives to reveal that information in the first place. Thus, an IO only leads I to punish A's defections—thus reducing these defections in the first place—for states that E wants to disclose intelligence about. These claims lead to two testable hypotheses, which are summarized in Table 1.

- *H1*: The greater the intelligence reception and protection capabilities of an IO, the more intelligence an informed state discloses to it about states it is not friendly with.
- *H2*: The greater the intelligence reception and protection capabilities of an IO, the fewer violations of international law that occur among states that are not friendly with the informed state.

Nuclear Intelligence and the IAEA

We apply these insights to the nuclear domain both for empirical tractability and due to the substantive importance of nuclear nonproliferation. States routinely gather intelligence about other states' nuclear programs by tracking overt and clandestine nuclear facilities, estimating stockpiled uranium or plutonium, and monitoring trade in sensitive technologies (Richelson 2007, 11–13). Such intelligence directly bears on whether a government is in compliance with its formal treaty commitments under the Nuclear Nonproliferation Treaty (NPT) as well as the broader norm against new nuclear weapons programs, in which a violation is defined as a nonnuclear state developing nuclear infrastructure and weaponization for military purposes. States may be tempted to share this information to facilitate diplomatic, economic, and other forms of pressure that might slow proliferation, particularly if the information pertains to non-allies (Fuhrmann and Kreps 2010). Yet sharing this intelligence risks revealing sensitive sources and methods, endangering the success of future intelligence collection efforts. For example, when Iraq found out that its nuclear activities had been tracked via its tapped cell phones, it stopped using them.[16] Syria, moreover, anticipated discovery via overhead imagery and attempted to conceal a clandestine nuclear facility from satellites.[17] Further, supplying intelligence conclusions without sources and methods raises credibility problems due to states' political biases, as highlighted by the United States' dubious intelligence claims about Iraq's program in 2002–3. As a consequence, states with sensitive intelligence often hold it back, making it more difficult to monitor the regime.

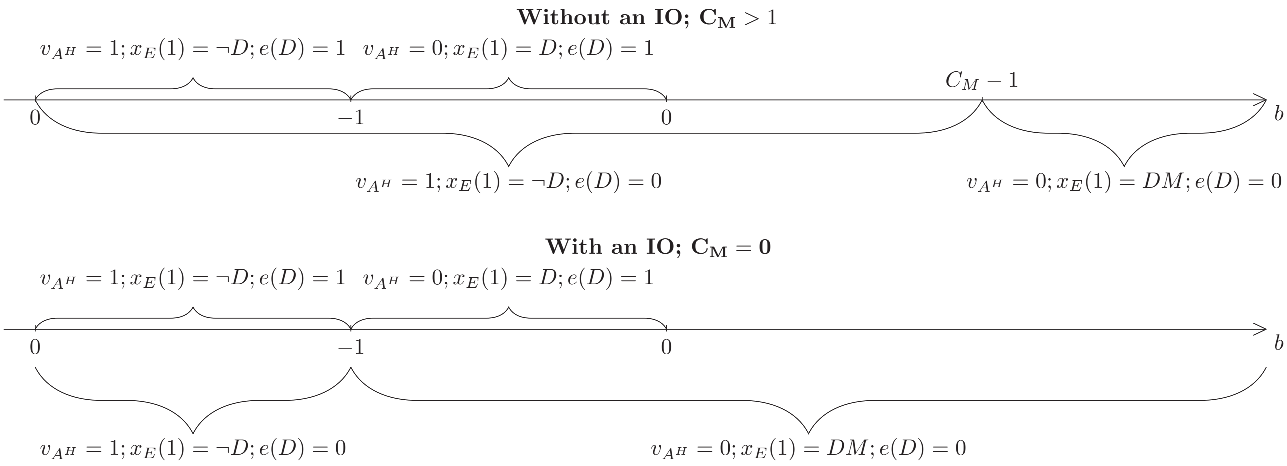
We argue that reforms that equipped the IAEA to receive, protect, and act on intelligence have helped it to alleviate these dilemmas. A core function of the agency is preventing the diversion of nuclear technology for nuclear weapons purposes through its safeguards agreements, which permit IAEA staff to visit states' declared nuclear facilities to confirm their records, monitor IAEA equipment, and verify declared uses of nuclear materials. Cases of suspected noncompliance are referred to the IAEA's Board of Governors, which may follow up with a state directly or refer the issue to the United Nations Security Council.[18] National intelligence could supplement the IAEA's information by highlighting discrepancies in declared activities, identifying undeclared but suspicious locations, and documenting efforts to evade detection (Acton 2014). However, while the IAEA had the technical

[16] See Aid (2009, 245).

[17] Amos Harel and Aluf Benn, *Haaretz*, March 23, 2018, "No Longer a Secret: How Israel Destroyed Syria's Nuclear Reactor." https://www.haaretz.com/world-news/MAGAZINE-no-longer-a-secret-how-israel-destroyed-syria-s-nuclear-reactor-1.5914407.

[18] If states have signed the Additional Protocol, the IAEA may also conduct more intensive inspections at undeclared sites as well.

FIGURE 2 Equilibria as a Function of Bias *b*


Without an IO; $\mathbf{C_M} > 1$
$v_{A^H} = 1; x_E(1) = \neg D; e(D) = 1$
$v_{A^H} = 0; x_E(1) = D; e(D) = 1$
$C_M - 1$
0
-1
0
b
$v_{A^H} = 1; x_E(1) = \neg D; e(D) = 0$
$v_{A^H} = 0; x_E(1) = DM; e(D) = 0$
With an IO; $\mathbf{C_M = 0}$
$v_{A^H} = 1; x_E(1) = \neg D; e(D) = 1$
$v_{A^H} = 0; x_E(1) = D; e(D) = 1$
0
-1
0
b
$v_{A^H} = 1; x_E(1) = \neg D; e(D) = 0$
$v_{A^H} = 0; x_E(1) = DM; e(D) = 0$

Note: Equilibria are described using the relevant part of players' strategies. In all equilibria, $v_A L = 0$, $x_E(0) = \neg D$, $e(\neg D) = 0$, and $e(DM) = 1$.

TABLE 1 Summary of Intelligence-Sharing Decisions

	IO Unable to Protect Intel	IO Able to Protect Intel
Intel about Friends	Rare Disclosures Violations Common	Rare Disclosures Violations Common
Intel about Non-Friends	Rare Disclosures Violations Common	Frequent Disclosures Violations Less Common

authority under Article VIII of the 1956 IAEA Statute to receive intelligence, it was not equipped to protect this information in practice.[19] After 1990, the agency argued that greater intelligence sharing was necessary for it to perform its duties (Ogilvie-White 2014, 325–26). It thus spearheaded reforms with its member states' agreement, which included clarification that IAEA staff would sign nondisclosure agreements, grant limited access to intelligence, implement procedures to address breaches of confidentiality, and secure cyber and physical infrastructure for sensitive information.[20] Moreover, informed states could provide their information via private briefings with the director general—whom powerful states ensure they trust during the selection process—and a select few staff members. Intelligence material could also be stored in the director general's office or other secure areas that are accessible only to thoroughly vetted and limited staff.[21] The implementation and credibility of these measures were buoyed by the IAEA's decades-long experience handling the sensitive data gathered in its routine inspections from member states.[22] The IAEA had also demonstrated its independence and neutrality in cases like South Africa, where some intelligence offers were refused to safeguard its integrity (Brown 2015, 108).

Observable Implications

The IAEA's reforms represent a shift in a prominent IO's ability to receive and protect intelligence—a key quantity of interest in our theoretical framework. We argue that if an IO cannot adequately protect E's sensitive information, E loses C_M and is reluctant to disclose it.[23]

[19] "Each member should make available such information as would, in the judgement of the member, be helpful to the Agency." Article VII, The Statute of the IAEA, 1956. On failure to exercise this authority, see Interview 3, June 20, 2017.

[20] Interview 4, June 22, 2017; Interview 6, June 20, 2017; Interview 43, February 18, 2017.

[21] Interview 14, February 9, 2017.

[22] Interview 43, February 18, 2017; Interview 14, February 9, 2017.

[23] For example, leaks from a United Nations peacekeeping mission in Somalia that received intelligence led attempts to cut off such

The institution's reforms revealed the IAEA's failure to detect Iraq's covert nuclear activities through its internal sources of information, highlighting the need for intelligence sharing.[24] Since the end of the Cold War damaged long-standing U.S.–Soviet cooperation to prevent nuclear proliferation among their respective allies, nonproliferation regime leaders such as the United States needed new ways to monitor and punish proliferation.[25] This may have been one reason Washington encouraged more effective IAEA monitoring via intelligence, as noted above.[26] Thus, 1991 constituted a "turning point" prompting a shift at the IAEA[27] because "information on clandestine programs was recognized as a gap that needed to be filled."[28]

We can thus use these reforms to assess variation over time in U.S. intelligence disclosures and other countries' nuclear activities. We focus on the United States because it possesses the most sophisticated intelligence collection capabilities of any state, has a long history of monitoring proliferation, provides more intelligence to the IAEA than other states, and offers archival records regarding intelligence and nuclear proliferation that are relatively accessible and consistent throughout the period we analyze.[29] Western European providers of intelligence tended to supplement and corroborate the basic claims from U.S. intelligence,[30] whereas other potential providers (e.g., Russia, China) did not value the mission of the IAEA in the same way and so rarely disclosed information despite the protections we detail.[31]

Our theoretical model shows that disclosure dilemmas only arise if an informed state has political reasons to disseminate information about another state's noncompliance, represented by the parameter b. An observable implication is that the United States should prefer to share intelligence about non-allies, since it is more threatened by a non-ally's nuclear arsenal than it is by an ally's. Furthermore, the United States possesses less costly tools for dissuading allies, such as trade ties, foreign aid flows, and other bilateral instruments (Miller 2014). We thus expect that the United States disclosed considerable intelligence to the IAEA about the nuclear activities of its non-allies in the post-1990 period, but rarely did so in the previous period. However, the United States should have rarely disclosed intelligence with the IAEA about the nuclear activities of its allies in either the post-1990 period or the previous period.

We further hypothesize that these intelligence disclosures should have reduced violations of international law (v) by improving the IAEA's ability to monitor suspicious nuclear sites and galvanizing multilateral enforcement efforts. Thus, facility closures for non-U.S. allies should have been more likely in the post-1990 period than in the previous period, whereas facilities for U.S. allies should have closed at a similar rate before and after the IAEA's reforms. This discussion is summarized in Table 2, which describes the theoretical quantity of interest along with the corresponding parameter from the model in the first column, and then indicates how it is operationalized and measured in the second column.

Empirical Analysis

We evaluate our hypotheses using all country-years in which a state pursued a nuclear military capability according to Bleek (2017).[32] A given state enters our data set if it pursued nuclear weapons. We then begin our analysis starting on the date from when they began exploring nuclear weapons or after the codification of the NPT in 1970, whichever came later, and end it when it either stopped pursuing nuclear weapons or successfully conducted a publicly acknowledged nuclear weapons test. If neither of these occurred, the analysis ends in 2015 due to the lack of reliable material from more recent years.[33] We include all states that meet these criteria regardless of

disclosures by the U.S. Congress (Wiebes 2003, 33–34). See the supporting information for additional details on UN peacekeeping and UNSCOM, a second IO in which leaks deterred disclosures.

[24] As one interviewee noted, "Iraq 1991 was a turning point" after which the IAEA began "using other sources of information." Interview 4, June 22, 2017. See the supporting information for further discussion.

[25] On collusion during the Cold War, see Coe and Vaynman (2015). On the decline of cooperation in the last decade, see Potter and Bidgood (2018).

[26] We thus argue that the end of the Cold War affected our outcomes primarily through the resulting IAEA reforms. As we discuss subsequently, other effects of the Cold War work against our findings, reducing the importance of intelligence sharing and increasing the feasibility of building nuclear facilities.

[27] Interview 3, June 20, 2017.

[28] Interview 4, June 22, 2017. Another interviewee stated that the impetus for reform "can be put exactly to 1991." Interview 45, September 19, 2017.

[29] Interview 7, May 22, 2017.

[30] Interview 2, June 20, 2017.

[31] Interview 7, May 22, 2017. Interview 2, June 20, 2017.

[32] We exclude the five states for which nuclear weapons are permitted under the NPT, which include the United States, China, Russia, Britain, and France.

[33] While India's test was not weaponized and thus was not coded as such by Bleek (2017), we include it because it publicly established

Table 2 Operationalizing Hypotheses in the Nuclear Domain

Variable	Concept (Parameter)	Operationalization (Data)
DV(H1)	Disclosure to IO (E chooses D, $\neg D$, DM)	U.S. intel sharing with IAEA (Archives/interviews)
DV(H2)	Violations (A chooses V, $\neg V$)	Nuclear facility closures (Fuhrmann 2015 data; archives/interviews)
IV(1)	IO can protect intel (C_M lower if E discloses to IO)	IAEA intel protection reforms (Post-1991 indicator)
IV(2)	$E - A$ political relationship (E receives b if I punishes rival)	Ally/non-ally (ATOP data set)

Table 3 Case Selection

	Pre-reforms (1970–90)	Post-reforms (1990–2015)
Ally	Israel (1970–90)	Israel (1991–2015)
	Pakistan (1972–90)	Pakistan (1991–98)
	Iran (1974–79)	
	South Korea (1970–81)	
	Taiwan (1970–76, 1987–88)	
	Brazil (1970–90)	
	Argentina (1978–90)	
Non-ally	Algeria (1983–90)	Algeria (1991)
	Iraq (1975–90)	Iraq (1991)
	Libya (1970–90)	Libya (1991–2003)
	Iran (1984–90)	Iran (1991–2015)
	North Korea (1970–90)	*North Korea (1991–2006)*
	South Africa (1970–90)	South Africa (1991)
	India (1970–74)	Syria (2000–7)

their NPT status, as the IAEA often had facility-specific safeguards agreements with non-NPT states that provided it with a monitoring role.[34] The resulting cases appear in Table 3; North Korea is italicized because we describe this case in full in the main text, whereas the others appear in the supporting information due to space constraints.[35]

To assess our first hypothesis, we collected new data on U.S. intelligence sharing with the IAEA from memoirs (e.g., ElBaradei 2011), public IAEA reports, archival material, secondary accounts of the IAEA and of specific states' programs, and interviews. Specifically, we interviewed a convenience sample of 33 experts, including a former IAEA director general, high-level IAEA staff, and current and former U.S. government officials.[36] While we record and describe each instance of this sharing in the case studies, our descriptive analysis codes this

India's mastery of the fundamentals for a nuclear weapon. See page 7 of the supporting information for details. We include "edge cases," which fell on the borderline between "explorers" and "pursuers," and include Taiwan (Bleek 2017, 40), Algeria (Bleek 2017, 45), and Argentina (Karp and Carasales 2000). Although the latter was not coded as an edge case by Bleek (2017), Karp and Carasales (2000) note that "powerful forces in the country clearly sought to develop and preserve the option."

[34] As we describe in the case study analyses, the IAEA maintained a significant relationship with many non-NPT states, including Pakistan, India, Israel, Argentina, and Brazil. Even without a significant formal relationship, the IAEA can still vet the validity of intelligence-based claims and allow the international community to follow up through ad hoc mechanisms, as in Iraq in the 1990s.

[35] We omit Yugoslavia due to the absence of declassified information about its activities and Ukraine because it inherited a publicly acknowledged nuclear program.

[36] We attempted to reach all IAEA officials for whom we were able to locate contact information, and interviewed those who would agree to do so.

dependent variable on a scale of 0 to 2, in which 0 indicates no sharing, 1 indicates few (1–3) instances of sharing, and 2 indicates more than three instances of sharing.

To evaluate our second hypothesis, we gathered new data on the number of uranium enrichment and plutonium reprocessing (ENR) plants either in construction or operation in a given country. While only military applications ("for explosive purposes") are technically prohibited by the NPT, the purpose of ENR facilities is usually ambiguous because of their potential for both civilian and military applications (Hoedl 2016, 66–67). Because of this dual use, they often indicate a state's potential interest in exploring or pursuing nuclear weapons. Operationalizing our dependent variable in this way is a hard test for our theory since closing a facility constitutes a particularly costly step to take. We thus begin with the list of ENR facilities from Fuhrmann and Tkach (2015) and then code whether the IAEA played a role in their closure using both primary and secondary sources, as described in the case studies. To be clear, the IAEA alone often does not have a decisive effect on these decisions. Instead, we expect intelligence to help the IAEA clarify the existence of suspicious activities and document evasion through site visits and the application of its technical expertise to intelligence-based allegations. Note that these data do not include facilities that were never built in the first place due to fears of an intelligence-enabled IAEA, adding to the difficulty of our empirical test.

Finally, we include two key independent variables. First, we add an indicator of the period before the IAEA's 1991 reforms. Second, we include an indicator of whether a given state is a U.S. ally according to Leeds (2005).[37] Note that we code Israel as a U.S. ally given its long-standing and well-recognized status as such (Bar-Siman-Tov 1998). We also code Taiwan as an ally after its formal alliance with the United States ended in 1980 because although the relationship became informal after the United States normalized relations with China, the two countries remained allies in practice.[38] For each case, we also describe the intelligence that the United States possessed in order to establish that it was capable of sharing such information with the IAEA.[39]

Aggregate Results

Table 4 summarizes our findings, listing a separate entry for each country before and after the IAEA's reforms so that most countries are shown twice. After coding our key independent and dependent variables, the table indicates whether each case demonstrated "strong," "moderate," or "weak" support for the theory. *Support for H1* was coded as "strong" if the United States shared intelligence with the IAEA more than three times when the theory would expect it to do so frequently (i.e., during the post-reform period for non-allies), and if no sharing occurred otherwise. It was coded as "weak" if more than three instances of sharing occurred when we anticipate sharing only rarely, or if no sharing occurred when we expect to find frequent sharing. Finally, it was coded as "moderate" if it was shared one to three times in either period. *Support for H2* was coded as "strong" if the IAEA's monitoring was improved by shared intelligence and its activities facilitated the closure of suspect facilities, and the reverse in periods before intelligence reforms. It was coded as "weak" if the IAEA's monitoring played a key role in closing facilities when our theory expects this to occur rarely, and when it did not do so otherwise. Lastly, we coded it as "moderate" when intelligence and the IAEA's monitoring played a minor role in the outcome.

Overall, we find that U.S. intelligence sharing was common after the reforms but not before, conditional on the ally status of the proliferating state. The provision of intelligence to the IAEA regarding U.S. allies in the pre-reform period was rare, as we find no evidence of sharing about Israel and South Korea and few instances about Pakistan and Taiwan. In the post-reform period, this pattern continued, as we did not code any instances of sharing about these four states. We also found few instances of intelligence provision in the pre-reform period (e.g., Brazil and North Korea). Yet after the IAEA reformed, intelligence was regularly supplied for each non-ally except for Algeria and Libya. For these two states, several instances of sharing occurred, representing a moderate increase from the pre-reform period.

Moreover, we find no evidence that U.S. intelligence sharing with the IAEA impacted its scrutiny of a U.S. ally's ENR facility and any decision to pause or shut down the facility. Nondisclosure and the resulting poor information at the IAEA were an enabling condition for nuclear progress for U.S. allies like Pakistan and Israel. Allies in the pre-reform period that made changes to ENR activities (e.g., South Korea and Taiwan) did so as a result of direct pressure from the United States rather than the IAEA's scrutiny, whereas all of the non-allies' programs were influenced at least somewhat by intelligence sharing

[37] Whether the alliance is formal or informal does not affect our results, as none of the non-allies that we consider were informal allies of the United States.

[38] Our results do not change if we instead code Taiwan as a non-ally after 1980.

[39] We note that this does not constitute a natural experiment since these changes were not made at random (e.g., Carnegie and Christoph 2017).

TABLE 4 Summary of Cases

Proliferator	Ally?	Pre-Reform?	Intel Shared (0-2)	# Facilities Closed	Support for H1	Support for H2
Algeria	No	Yes	0	0	Strong	Strong
Algeria	No	No	1	1[b]	Moderate	Moderate
Argentina	Yes	Yes	0	0	Strong	Strong
Brazil	Yes	Yes	1	0	Moderate	Strong
India	No	Yes	0	0	Strong	Strong
Iraq	No	Yes	0	0	Strong	Strong
Iraq	No	No	2	5	Strong	Strong
Iran[a]	No	Yes	0	0	Strong	Strong
Iran	No	No	2	5	Strong	Strong
Israel	Yes	Yes	0	0	Strong	Strong
Israel	Yes	No	0	0	Strong	Strong
Libya	No	Yes	0	0	Strong	Strong
Libya	No	No	1	1	Moderate	Strong
North Korea	No	Yes	1	0	Moderate	Strong
North Korea	No	No	2	1	Strong	Strong
Pakistan	Yes	Yes	1	0	Moderate	Strong
Pakistan	Yes	No	0	0	Strong	Strong
South Africa	No	Yes	0	0	Strong	Strong
South Africa	No	No	2	1	Strong	Strong
South Korea	Yes	Yes	0	0	Strong	Strong
Syria	No	No	1	0	Moderate	Moderate
Taiwan	Yes	Yes	1	0	Moderate	Strong

Note: [a] Iran was an ally from 1974 to 1979. [b] Algeria's facility was not technically closed, but it accepted IAEA safeguards, as described in the supporting information.

with the IAEA in the post-reform period. To get a sense of how intelligence assisted with the rollback of individual facilities, Table 5 illustrates our mechanism, giving examples of ENR facilities in non-ally states whose closure was likely due to intelligence-enabled IAEA scrutiny. Our case study narratives address these facilities in more detail, and a complete table of facilities appears in Table 1 of the supporting information.

We also consider whether other effects of the Cold War's end might explain the increased intelligence sharing and reduced success in developing ENR facilities among non-U.S. allies.[40] However, our mechanism-specific evidence casts doubt on the importance of slow-moving, structural changes. First, our case study evidence makes it clear that *interest* in nuclear proliferation did not change due to the end of bipolarity. North Korea, described subsequently, maintained the view that a nuclear deterrent was critical to its survival during and after the Cold War. The closure of its ENR facilities did not occur as a result of a changing threat and a resulting disinterest in a nuclear arsenal. Second, our case studies feature evidence that specific intelligence disclosure decisions about particular facilities enabled greater IAEA scrutiny, boosted multilateral pressure, and dissuaded proliferators from continuing such facilities. In pre–Cold War cases like North Korea, moreover, we show that the United States' reluctance to disclose intelligence due to sources and methods concerns insulated it from scrutiny. Together these suggest that larger changes in threat and geopolitics do not explain the intelligence disclosure and facility-related patterns that we identify.

Moreover, other changes triggered by and coinciding with the end of the Cold War work against our findings. The end of bipolarity led to significant cuts to defense and intelligence budgets, including that of the United States (Baker and Williamson 2000). The value of clandestine information also declined as open-source information became more useful for identifying proliferation activities, including commercial satellite imagery (Warner 2014). Thus, in the absence of changes at the IAEA, the likely pattern would have been less frequent intelligence disclosures after the Cold War. Moreover, the end of bipolarity and

[40] We discuss additional possibilities in the supporting information.

TABLE 5 Intelligence and Closed Facilities: Illustrative Examples

Country	Enrichment/Reprocessing Facility	Intelligence Role
Iran	Tehran Nuclear Research Center (Reprocessing)	U.S. intel sharing stepped up in 1992 and 1993; IAEA visit in February 1992 includes three new sites based on Western intel
	Plasma Physics Laboratories in Tehran	IAEA site visits and intelligence leads encourage Iran to move centrifuge research to less well-known location (Kalaye) in mid-1990s
	Kalaye Electric Company	U.S. satellite photos reveal site and evidence of Iranian cover-up; initial move to Kalaye in part to avoid IAEA scrutiny at TNRC
	Lashkar Ab'ad	Western intel supports IAEA inspections of laser enrichment in 2003; Iran ends work as scrutiny intensifies; U.S. intel sharing regularly in 2003/2004
Iraq	Rashdiya Building 22	IAEA fails to detect enrichment work at Rashdiya in 1991 visit; intel sharing and defector information eventually help IAEA confirm in mid-1990s
	Al Tuwaitha Hot Cell	U.S. intel sharing shows cover-up near Tuwaitha and enables IAEA follow-up visit that helps secure permanent closure
	Al Tarmiya	U.S. intel identified Tarmiya as likely facility before war but did not share; overhead photos shared in 1991 help IAEA identify enrichment work here
Libya	Tajoura Enrichment Facility	U.S. gets new intel on Russian scientists and AQ Khan around 1991/1992 IAEA visits in February 1992; extent of intel sharing with IAEA unclear but likely
North Korea	Radio Chemical Laboratory (Yongbyon)	U.S. imagery key to initial IAEA scrutiny; more U.S. intel sharing helps IAEA visits (1992–1993) that facilitate diplomatic deal ending Yongbyon work
South Africa	Valindaba Z - Plant	U.S. intel briefing to IAEA assists in verification of dismantlement; IAEA verifies shutdown and reported amounts of low enriched uranium

Note: Sources are provided in the supporting information.

independent technological advances made the technical know-how in enrichment and reprocessing more accessible. Without more effective scrutiny by the IAEA, we thus expect that the proliferation of ENR facilities would have been more, not less, common in the post–Cold War era.

Case Studies

We present the case of North Korea here, and the remaining cases appear in the supporting information due to space constraints. Because the North Korean case straddles the pre- and post-reform period, we can analyze within-case variation, allowing for more fine-grained inferences about our hypotheses and the theory's causal mechanisms. We first provide an overview of North Korea's program and describe the United States' knowledge about it. We then show that the United States withheld intelligence before 1991 but provided considerable information to the IAEA afterward in response to its reforms, in accordance with our first hypothesis. We outline how this new information improved the IAEA's ability to monitor the program and helped to slow North Korea's nuclear progress, in line with our second hypothesis.

North Korea, 1970–2006

While North Korea's civilian nuclear program started in the 1950s, its weapons development began in earnest in the 1980s. North Korea focused on plutonium reprocessing and reportedly began high explosives testing for an implosion device around 1989 (Wit, Poneman, and Gallucci 2004, 6). The IAEA's detection of discrepancies in Pyongyang's reports led to intense multilateral pressure in 1993. Rather than submit to special inspections, North Korea stated its intention to withdraw from the NPT, which led to intense diplomatic negotiations and an agreement ("Agreed Framework") in 1994 to freeze North Korea's reprocessing activities.[41] Disputes over the agreement's implementation and revelations of undeclared uranium enrichment led to its collapse in 2003 (Chinoy 2010, 117–26) and North Korea's nuclear warhead test in 2006.

The United States possessed considerable intelligence about North Korea's program. Though Soviet leaders did not convince North Korea to join the NPT until 1985 (Oberdorfer and Carlin 2013, 198) and to complete a comprehensive safeguards agreement until 1992,[42] the United States had monitored North Korea's program continuously since 1950. Overhead imagery revealed the construction of a new large-scale power reactor at Yongbyon in the early 1980s, a reprocessing facility and high explosives testing in the late 1980s (Wit, Poneman, and Gallucci 2004, 1–6), and concealed waste storage in the early 1990s. Defectors also contributed to the United States' receipt of a "steady trickle" of intelligence (Richelson 2007, 358).

Our first hypothesis is that the United States withheld its intelligence from the IAEA before 1991 and disclosed it afterward, as the United States and North Korea have been adversaries since the early 1950s (Leeds 2005). Indeed, prior to 1991, the United States "held much back" and was especially reluctant to release "imagery from newer systems." It occasionally provided outdated satellite images of the nuclear facilities (Richelson 2007, 519) but preferred to limit any sharing of intelligence to North Korea's patrons (e.g., the Soviet Union and China; Oberdorfer and Carlin 2013, 198–99). Leaks were a serious concern, as intelligence officials "feared the unauthorized disclosure of its intelligence sources or methods."[43] Several interviewees also noted that intelligence disclosures to the IAEA were very rare in the 1980s.[44]

However, the first Iraq War and reforms at the IAEA in 1991 prompted a shift in behavior, as "the IAEA under Hans Blix underwent an upheaval in personnel and a sea change in attitudes" and "established the right to accept intelligence information supplied by the United States and other member states in its investigations" (Oberdorfer and Carlin 2013, 209). Beginning in 1991, the United States thus began to brief "a small group of IAEA experts" on North Korea using "spy satellite photos and other highly sensitive intelligence" that it had previously kept off-limits.[45] An IAEA official stated that "we had never received stuff like this before" and highlighted that "the pictures gave us prima facie evidence that the facilities are nuclear-waste related."[46] Our interviewees noted this shift as well. One stated that it was only "after 1991" that

[41] See IAEA, "IAEA and DPRK: Chronology of Key Events." https://www.iaea.org/newscenter/focus/dprk/chronology-of-key-events.

[42] See Wit, Poneman, and Gallucci (2004, 13). North Korea, the Soviet Union, and the IAEA had a facility-specific safeguards agreement in 1977 for the Soviet-built research reactor. See IAEA, "IAEA and DPRK: Chronology of Key Events." https://www.iaea.org/newscenter/focus/dprk/chronology-of-key-events.

[43] R. Jeffrey Smith, "N. Korea and the Bomb: High-Tech Hide-and-Seek," *Washington Post*, April 27, 1993. Smith additionally reports that the intelligence community also sought exchanges of information with the IAEA.

[44] Interview 3, June 20, 2017; Interview 10, February 16, 2017.

[45] Smith, "High-Tech Hide-and-Seek."

[46] Smith, "High-Tech Hide-and-Seek."

the United States "provided intelligence about a particular [North Korean] site" despite tracking activity for some years prior,[47] while another highlighted the U.S. intelligence briefings about North Korea in 1992.[48] Oberdorfer and Carlin (2013, 209–10) explain:

> Starting in September 1991, in the wake of the Gulf War, the United States began supplying intelligence information to Blix and senior aides at his Vienna headquarters, eventually including the services of its incomparably sophisticated national laboratories and supplied photos from US spy satellites that were rarely shown to outsiders. Armed for the first time with extensive independent information about the nuclear programs that they were checking, the IAEA's leaders and its corps of international inspectors were determined not to be hoodwinked or embarrassed again. *North Korea became the first test case of their new capabilities and* attitudes. . . . Taking no chances that the IAEA chief would miss something of importance, US officials had provided intelligence briefings for Blix and his top aides in September 1991, March 1992, and on May 7, immediately before his departure.[49]

Consistent with our first hypothesis, institutional reforms played an important role in this shift to more intelligence disclosures. The above quote, for example, refers to North Korea in the early 1990s as a "test case" of new IAEA techniques, which suggests that the institution's new authority and procedures regarding intelligence were seen by Washington as a tool to potentially improve its performance. Moreover, one interviewee noted that, as of the late 1990s, intelligence was "routinely provided." Asked why American leaders trusted the IAEA with sensitive details that they were not comfortable releasing publicly, the interviewee replied that the IAEA had developed "a rigid protocol" by the 1990s for handling information within the Secretariat and that, over time, it "built a track record." Confidence in its ability to handle intelligence was, by then, "very high."[50] A second interviewee also pointed to the IAEA's track record and contrasted early, isolated instances of sharing in the late 1980s with that in the 1990s, noting that "over the years, the [intelligence community] has become much, much more comfortable with sharing information with someone like the IAEA." One indicator of that comfort, the interviewee noted, was the existence of far fewer internal battles within the U.S. government about the risks to sources and methods when intelligence was disclosed to the IAEA in later years.[51]

Our second hypothesis expects this intelligence to have assisted the IAEA with its scrutiny of suspect facilities, and indeed, it was useful for both identifying hidden sites and documenting attempted cover-ups. First, intelligence helped to identify hard-to-find, undeclared ENR facilities, specifically the Radio Chemical Laboratory (Yongbyon) reprocessing facility and its related buildings. Shared intelligence revealed North Korean workers constructing and concealing a new storage site (Richelson 2007) and a hidden floor in a nuclear waste facility (Oberdorfer and Carlin 2013, 214), and it helped to prepare IAEA inspectors for visits of these sites. Prior to the IAEA's first visit in May 1992, a Central Intelligence Agency briefing of the IAEA director general and his staff included "a 'virtual reality' tour of Yongbyon using advanced computer modeling based on aerial photographs" (Oberdorfer and Carlin 2013, 210). Second, U.S. disclosures provided unique evidence of North Korea's active deception, including overhead imagery showing North Koreans "moving equipment" away from scrutinized sites, "demolishing freshly built walls," and an "old site being buried under dirt and dozens of hastily planted shrubs and trees."[52] Successive IAEA visits validated these suspicions and added credibility to the claim that Pyongyang was hiding its efforts to build a bomb. These milestones were critical in winning approval for IAEA visits from other member states as well as the diplomatic negotiations that produced the Agreed Framework (Dembinski 1995, 33–34).[53] Overhead imagery given to the IAEA by the United States "marked a turning point in the investigation" (Acton 2014, 346), and intelligence presented in a closed-door, "highly classified" briefing to other Board of Governor members was described by one interviewee as having a "huge impact" on subsequent voting in 1993.[54]

Developments in the 2000s also provide support for our hypotheses. First, the collapse of the Agreed Framework in 2003 was partly due to the United States choosing

[47] Interview 41, August 7, 2017.

[48] Interview 2, June 20, 2017.

[49] Emphasis added.

[50] Interview 7, May 22, 2017.

[51] Interview 10, February 16, 2017.

[52] The first two quoted phrases are from Richelson (2007, 519). The third is from Smith, "High-Tech Hide-and-Seek."

[53] One of the terms of the Agreed Framework was ending work on Yongbyon, though the Isotope Production Laboratory remained operational.

[54] Interview 3, June 20, 2017. The interviewee noted other unusual security measures, such as the removal of non-Board member states, nonessential IAEA staffers, and the continued high-level classification of the meeting records two decades later.

to bypass the IAEA with its intelligence, instead making unverified claims directly to North Korea about alleged undeclared uranium enrichment.[55] Consistent with the theory, these unilateral claims were met with considerable skepticism.[56] One reason may have been a reduced ability for the IAEA to validate or supplement intelligence-based claims with additional monitoring visits. Second, our interviews featured evidence that the United States continues to update the IAEA via intelligence briefings in case the IAEA is permitted to resume inspections.[57] Even after North Korea's 2006 nuclear test, the IAEA prepared to resume inspections by staying updated via open-source information and intelligence disclosures.[58]

Conclusion

This article identifies a pervasive problem in international relations, showing that states with intelligence about international rule violations often decline to share it due to fears of damaging their future intelligence collection abilities, but that properly equipped IOs can help to mitigate this issue when states have political incentives to provide it. Our results have a number of surprising implications. First, while it is intuitive that the United States would want to harm its adversaries and protect its allies, our finding that the United States does so by disclosing intelligence to an international organization—rather than going public with its conclusions or sharing them quietly with its allies—is striking. Given common concerns about sovereignty and information security, it is surprising that the United States systematically provides intelligence to an IO like the IAEA at all. Second, our demonstration that increased secrecy improved a prominent IO's effectiveness is not anticipated by existing studies in this area, which would expect increased institutional transparency to do so. Our results thus complicate current understandings of the determinants of IO effectiveness and the mechanisms by which IOs improve cooperation in international politics.

While we test our argument in the nuclear realm, the concepts and theory we develop have broad applicability in international relations. Countries with sophisticated intelligence capabilities often possess unique information that pertains to the guilt or innocence of individuals suspected of genocide, crimes against humanity, or other war crimes. Sharing this information can promote international justice, satisfying international and domestic pressures to "do something" about the situation, but may reveal the sources and methods used to obtain it. In the international trade and investment arenas, states that are accused of violating the rules can be tempted to reveal sensitive firm-level details, such as contract terms or trade secrets, to bolster their case for innocence. Yet supplying commercially sensitive information can allow competitors to damage the disclosing firms' market competitiveness. Similar trade-offs exist regarding commercially sensitive information relevant to international financial regulations, intelligence relevant to peacekeeping missions, and other areas.

Moreover, the relevance of disclosure dilemmas and use of IOs to address them is further underscored by the diversity of IOs that are designed to protect sensitive information. For example, the Organization for the Prohibition of Chemical Weapons maintains the "rigid confidentiality" of sensitive information about chemical industries; the International Monetary Fund's Financial Sector Assessment Program uses a three-tiered document classification system to protect states' financial sector health; the 1989 Montreal Protocol that regulates the emission of chlorofluorocarbons obscures commercially valuable information; and the World Trade Organization protects trade secrets and other firm-level data from public exposure during trade disputes.[59] However, although our theory is broadly applicable, it does not apply to every IO. Disclosure dilemmas only arise when an informed state would both obtain a benefit and incur a cost from revealing its information. Without a benefit, the state would simply keep its information private, and without a cost, the state would reveal it regardless of whether an IO was available. In domains in which disclosure dilemmas do not arise, IOs may instead reduce transaction costs, serve as clubs of states with similar interests, solve political hold-up problems, or play other roles (Carnegie 2015;

[55] On the IAEA's exclusion from this confrontation and subsequent diplomacy, see ElBaradei (2011, 90–94). Details on the role of intelligence are in Chinoy (2010, 115–18).

[56] Russia, for example, complained that the United States' allegations were only "based on one source." See Michael Wines, "Briefed Further on North Korea, the Russians, Openly Skeptical Before, Are Silent," *New York Times*, October 23, 2002. https://www.nytimes.com/2002/10/23/world/threats-responses-moscow-briefed-further-north-korea-russians-openly-skeptical.html. Similarly, "doubts began to surface in the press about whether North Korea was constructing a bricks-and-mortar facility." See Jeffrey Lewis, "How A.Q. Khan Helped Distort America's DPRK Policy," *38North*, March 29, 2010. https://www.38north.org/2010/03/how-a-q-khan-helped-distort-america's-dprk-policy/.

[57] Interview 9, March 31, 2017.

[58] Interview 2, June 20, 2017.

[59] See, respectively, Article VIII of the Chemical Weapons Convention; the Articles of Agreement of the International Monetary Fund, Article V, Section 2(B); Chayes and Chayes (1998, 165); and Grando (2009, 276). See the supporting information.

Davis and Wilf 2017). Moreover, even when these dilemmas are present, IOs must be properly equipped to solve them, as discussed previously.

Our results have many theoretical and normative implications. In our theory, IOs address disclosure dilemmas through institutional secrecy. Such secrecy is in tension with expectations of transparency and accountability in global governance.[60] Advocates of transparency norms argue that transparency can enhance international institutions' legitimacy in an era in which these organizations often face intense scrutiny and populist scorn even in established Western democracies.[61] This article highlights a potential trade-off between legitimacy and performance: Improving compliance-related activities by including intelligence can create new sources of opacity in IOs. Recognizing that IOs can solve multiple types of information problems, which require different solutions, thus enables scholars and practitioners to more fully understand the challenges and complications of efforts to make IOs accountable to a wider audience. Moreover, the way in which sensitive information provides a subtle source of power adds to this challenge. While the extant literature often argues that states derive power from holding key positions within IOs, funding these bodies, and engaging in bribery,[62] we show that states with intelligence can choose to share or withhold their information based on political interests. This shapes whose suspected noncompliance gets scrutinized, caught, and punished. Even when an IO—such as the IAEA—carefully protects its neutrality, the integration of intelligence provides some states with influence over the distribution and quality of monitoring. These implications for transparency and power, and their manifestation in other empirical domains and IOs, constitute promising topics for future research.

References

Acton, James M. 2014. "International Verification and Intelligence." *Intelligence and National Security* 29(3): 341–56.

Aid, Matthew M. 2009. *The Secret Sentry: The Untold History of the National Security Agency*. New York: Bloomsbury.

Baker, John C., and Ray A. Williamson. 2000. "The Implications of Emerging Satellite Information Technologies for Global Transparency and International Security." In *Power and Conflict in the Age of Transparency*, ed. Bernard I. Finel and Kristin M. Lord. New York, NY: Springer, 221–55.

Bar-Siman-Tov, Yaacov. 1998. "The United States and Israel since 1948: a 'special relationship'?" *Diplomatic History* 22(2): 231–62.

Bleek, Philip C. 2017. "When Did (and Didn't) States Proliferate? Chronicling the Spread of Nuclear Weapons" Discussion Paper. Cambridge, MA: Project on Managing the Atom, Belfer Center for Science and International Affairs, Harvard University and Monterey, CA: James Martin Center for Nonproliferation Studies, Middlebury Institute of International Studies, June 2017. https://www.belfercenter.org/sites/default/files/files/publication/When%20Did%20%28and%20Didn%27t%29%20States%20Proliferate%3F_1.pdf.

Brown, Robert L. 2015. *Nuclear Authority: The IAEA and the Absolute Weapon*. Washington, DC: Georgetown University Press.

Carment, David, and Martin Rudner, eds. 2007. *Peacekeeping Intelligence: New Players, Extended Boundaries*. New York: Routledge.

Carnegie, Allison. 2015. *Power plays: how international institutions reshape coercive diplomacy*. Cambridge: Cambridge University Press.

Carnegie, Allison, and Austin Carson. 2018. "The Spotlight's Harsh Glare: Rethinking Publicity and International Order." *International Organization* 72(3): 627–57.

Carnegie, Allison, and Austin Carson. 2019. "Reckless Rhetoric? Compliance Pessimism and International Order in the Age of Trump." *The Journal of Politics* 81(2).

Carnegie, Allison, and Mikulaschek Christoph. 2017. "The Promise of Peacekeeping: Protecting Civilians in Civil Wars." *SSRN*. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2909822

Carson, Austin. 2018. *Secret Wars: Covert Conflict in International Politics*. Princeton, NJ: Princeton University Press.

Chayes, Abram, and Antonia Handler Chayes. 1998. *The New Sovereignty*. Cambridge, MA: Harvard University Press.

Chesterman, Simon. 2006. "Shared Secrets: Intelligence and Collective Security." *Lowy Institute Paper* (10).

Chinoy, Mike. 2010. *Meltdown: The Inside Story of the North Korean Nuclear Crisis*. New York: St. Martin's Press.

Coe, Andrew J., and Jane Vaynman. 2015. "Collusion and the Nuclear Nonproliferation Regime." *Journal of Politics* 77(4): 983–97.

Coe, Andrew J., and Jane Vaynman. 2018. "The Tragedy of Arming." *Unpublished manuscript*.

Dai, Xinyuan. 2002. "Information Systems in Treaty Regimes." *World Politics* 54(4): 405–36.

Davis, Christina L., and Meredith Wilf. 2017. "Joining the Club: Accession to the GATT/WTO." *Journal of Politics* 79(3): 964–78.

Dembinski, Matthias. 1995. "North Korea, IAEA Special Inspections, and the Future of the Nonproliferation Regime." *The Nonproliferation Review* 2(2): 31–39.

Dreher, Axel, Peter Nunnenkamp, and Rainer Thiele. 2008. "Does U.S. Aid Buy UN General Assembly Votes? A Disaggregated Analysis." *Public Choice* 136(1–2): 139–64.

ElBaradei, Mohamed. 2011. *The Age of Deception: Nuclear Diplomacy in Treacherous Times*. New York: Macmillan.

[60] However, secrecy can sometimes prevent escalation (Carson 2018).

[61] For instance, see Posner (2017).

[62] For instance, see Dreher, Nunnenkamp, and Thiele (2008).

Fuhrmann, Matthew, and Sarah E. Kreps. 2010. "Targeting Nuclear Programs in War and Peace: A Quantitative Empirical Analysis, 1941–2000." *Journal of Conflict Resolution* 54(6): 831–59.

Fuhrmann, Matthew, and Benjamin Tkach. 2015. "Almost Nuclear: Introducing the Nuclear Latency Dataset." *Conflict Management and Peace Science* 32(4): 443–61.

Grando, Michelle T. 2009. *Evidence, Proof, and Fact-Finding in WTO Dispute Settlement*. Oxford: Oxford University Press.

Hoedl, Seth. 2016. "Ensuring Peaceful Use via International Licensing of the Nuclear Fuel Cycle." In *Nuclear Non-Proliferation in International Law, Volume III*, ed. Jonathan L. Black-Branch and Dieter Fleck. The Hague, The Netherlands: Springer, 63–112.

Karp, Aaron, and Julio Carasales. 2000. "Argentina and the Bomb." *The Nonproliferation Review* 7(1): 189–90. https://doi.org/10.1080/10736700008436805

Keohane, Robert O. 1984. *After Hegemony: Cooperation and Discord in the World Political Economy*. Princeton, NJ: Princeton University Press.

Leeds, Brett Ashley. 2005. "Alliance Treaty Obligations and Provisions (ATOP)." Rice University, Department of Political Science, Houston, Texas.

Lindsey, David. 2015. "Military Strategy, Private Information, and War." *International Studies Quarterly* 59(4): 629–40.

Miller, Nicholas L. 2014. "The Secret Success of Nonproliferation Sanctions." *International Organization* 68(4): 913–44.

Oberdorfer, Don, and Robert Carlin. 2013. *The Two Koreas: A Contemporary History*. London: Hachette.

Ogilvie-White, Tanya. 2014. "The IAEA and the International Politics of Nuclear Intelligence." *Intelligence and National Security* 29(3): 323–40.

Posner, Eric A. 2017. "Liberal Internationalism and the Populist Backlash." *Arizona State Law Journal* 49(Special Issue): 795–819.

Potter, William, and Sarah Bidgood. 2018. "Once and Future Partners: The United States, Russia and Nuclear Non-Proliferation." *IISS Adelphi Series*. https://www.iiss.org/publications/adelphi/2018/once-and-future-partners.

Richelson, Jeffrey. 2007. *Spying on the Bomb: American Nuclear Intelligence from Nazi Germany to Iran and North Korea*. New York: Norton.

Slantchev, Branislav L. 2010. "Feigning Weakness." *International Organization* 64(3): 357–88.

Thompson, Alexander. 2006. "Coercion through IOs: The Security Council and the Logic of Information Transmission." *International Organization* 60(1): 1–34.

Voeten, Erik. 2005. "The Political Origins of the UN Security Council's Ability to Legitimize the Use of Force." *International Organization* 59(3): 527–57.

Walsh, James Igoe. 2010. *The International Politics of Intelligence Sharing*. New York: Columbia University Press.

Warner, Michael. 2014. *The Rise and Fall of Intelligence: An International Security History*. Washington, DC: Georgetown University Press.

Wiebes, Cees. 2003. *Intelligence and the War in Bosnia*, 1992–1995. Vol. 1. Münster: LIT Verlag.

Wit, Joel S., Daniel B. Poneman, and Robert L. Gallucci. 2004. *Going Critical: The First North Korean Nuclear Crisis*. Washington, DC: Brookings Institution Press.

Supporting Information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

- **Case Study Appendix**
- **Model Proofs**
- **Alternative Explanations**
- **Why IAEA Reforms?**
- **Generalizability**
- **Data Collection on IOs**